

## **Commentary on “The Potential of the Internet for Music Perception Research: A Comment on Lab-Based Versus Web-Based Studies” by Honing & Ladinig**

ROGER A. KENDALL

*Music Cognition and Acoustics Laboratory  
Herb Alpert School of Music  
University of California, Los Angeles*

**ABSTRACT:** The use of web-based data collection raises fundamental issues impinging on reliability and validity of test results, as well as important ethical (and potentially legal) issues such as informed consent.

Submitted 2008 February 4; accepted 2008 February 4

**KEYWORDS:** *Internet-based experimenting, empirical, experimental method*

HONING and Ladinig (2008) bring forth an important topic regarding the specific issue of using the internet for data collection, and in the broadest sense of replacing data collection of music perception experiments with what they suggest is greater ecological validity with such an approach. My analysis and commentary will focus on responding to such claims, and emphasizing the issue of human subjects protocols for publication.

The allure of on-line data collection is clear in that more subjects can theoretically participate with automatic collection of information outside of the laboratory. As they suggest, at a cursory level, this value is clear in that ecological validity across many different subject backgrounds and stimulus presentation conditions appear to be provided by this approach. However, reliability is the counterpoint to validity. As validity in this context may potentially increase, reliability can decrease. Central to this point is the empirical investigation process itself. Both elements of reliability and validity are juxtaposed in relation to the problems and questions raised as the reason for doing a perceptual study in the first place. The questions and hypotheses of an investigation should ultimately control what approach is taken and not the reverse. For a full discussion of the relation between experimental variables in operational definition and reliability and validity, see Kerlinger (1999).

It is obvious that if the goal of an experiment is replication of results, increasing the ecological validity in the sense of myriad equipment at the subjects' on-line computing system, may detract or abandon a sense of reliability. Ultimately, the operational definitions of the study must without fail mirror the methods employed to provide an answer. If the experiment approaches the control required of psychoacoustical experiments, for example timbre ratings, information about whether the presentation is free field or via headphones is important, as is basic loudness and flat-response data about the transducer. How can these variables be determined on-line? One can imagine using specialized pre-test experimental data collection and real-time analysis, but this would extend the length of the experiment, likely compromising any advantage of the on-line approach. It would also require sophisticated algorithms for determining minimal sound reproduction conditions.

Remember that on-line browsers exist for portable systems such as cell phones and media players; variability of both visual and auditory parameters is vast in the on-line context. My current work with animation and musical pattern superpositions could not reasonably be presented on miniature screens in a commuter environment. The presupposition that experiments on-line are conducted in something similar to a home context with a 'standard' computer installation is increasingly compromised by the proliferation of portable systems.

Systematic and detailed methodological studies comparing multivariate stimulus and response conditions with laboratory-based and on-line approaches are required. Methodological investigations were once common in psychophysics (method of limits, neural quantum theory, method of constant stimuli, etc).

Such methodological experiments are increasingly rare in music perception, perhaps because scholarly publications consider new data more important than data arising from methodological comparison.

Honing and Ladinig present a single study (McGraw et al. 1999) with comparative analysis using on-line and laboratory experiments. The conclusion is not compelling, with the suggestion that the sample sizes achievable over the web obviate stimulus variability issues. This requires extensive proof. The further comments regarding perceptual based coding of sound, such as Mpeg4 (or Mpeg 3 or ATRAC 1 presume) seem out of place, unless one is concerned about bandwidth of transmission. Why would one stream video and audio on-line and not transmit the entire experiment intact for execution on the receiving device with the appropriate parameters? For example, surely timing data for rhythm perception would not be collected with streaming techniques in both directions since this could compromise accuracy at the mercy of packet-based transmissions, and the level of reliable bit rate transmission would have to be constantly monitored by additional software programming.

Honing and Ladinig suggest several steps to minimize the dropout rates (typically around 30-40%) of on-line based experiments. The audio and browser check suggestion lacks specificity. A query of the soundcard specifications and browser elements could be considered data collection and potential invasion of privacy. It is certainly possible to do such a check at the system level using Java. However, how the data are collected and utilized and made secure could be problematic (further discussion of human subjects issues follows). The suggestion of engaging instructions is also left without specificity. What makes instructions engaging other than clarity? The restriction of an experiment to 15 minutes or less at their suggestion may be based on their implementations, but it is certainly a straight-jacket. As I indicated earlier, *the experiment design must be at the service of the questions and hypotheses*. Designing the experiment to be engaging via a screen cast video and making the presentation ‘compelling’ *may* be completely at odds with the goal of the experiment.

Honing and Ladinig dismiss issues of “University Review Panels” and the oversight of “some high impact journals” in the thesis statement of the article at the end of the first paragraph. I strongly take issue with this. Journals and conferences have agreed that government guidelines and mandates for human subjects must be followed. For example, the Acoustical Society of America requires a presenter to digitally sign that such HS protocols have been followed in order to present at the conference. Within each country, guidelines and protocols are locally interpreted and evaluated by the research institution that hosts the experiment. My comments here are based on the UCLA Office for Protection of Research Subjects Policy Number 44 (2007) (<http://www.oprs.ucla.edu/human/documents/pdf/44.pdf>), which is in turn based on “Ethical and Legal Aspects of Human Subjects Research on the Internet” (<http://www.aaas.org/spp/sfml/projects/intres/report.pdf>), a highly detailed and documented US government workshop report (1999).

The overriding issue is that internet data collection and experimentation must follow all considerations and particulars that laboratory based studies must follow. Generally, an Institutional Review Board (IRB) must evaluate research requests on a cost/benefit basis considering factors such as risks of violation of privacy, voluntary participation, informed consent, and data confidentiality.

I suggest that web-based experiments require extraordinary steps to meet the high level of required confidentiality and security demanded, and that this practically eliminates the advantage of using an on-line approach. Legal issues have arisen regarding the ability of hackers to retrieve information for on-line sources; laboratory computers can be completely off-line. In situations where rewards (prizes or monetary compensation) are provided, often the experimental system would need to collect personal identification data (which must be kept randomized relative to experimental data). For a laboratory study, such data can be recorded on paper and transferred to a separate accounting office with no data connection to the experiment server. If all data are kept on-line, different server levels could be potentially exposed to hacker intervention. IRB analysis of the server security must be extensive and could result in considerable time for a project proposal to be approved.

Demographic survey data can easily be randomized so that connections of name, age, background need not be indexed to subject response data in both laboratory and web-based settings. However, a laboratory is usually within a sub-population context such as a university. As such, an important safeguard, that of exposing children without parental permission to an experiment, is often lost on web-based data collection. In the United States, additional protocols and legal requirements are specified regarding children (see Children’s Online Privacy Protection Act (COPAA) (<http://www.ftc.gov/ogc/coppa1.htm>)). It is arguable that demographic misinformation is constrained in a laboratory setting. In fact, on-line experiments can either be open to the general public or restricted to a sub-population group, such as

students or music majors. However, this again requires setting up connections between the experiment server and university registrar servers, something exceedingly unlikely to be allowed and very difficult at the level of programming and database information transfer. [1]

It is typical for the IRB to require a formal statement of the cost/benefit aspects of the experiment to society and the subject, outline any potential risks, and inform that subject that the experiment is voluntary and that they can withdraw without consequence at any time. This is easily accomplished with a check box on-line. My current experiments, which are certified IRB exempt, still require a two-page single-spaced consent form with seal and signatures in duplicate for each subject. It is possible that, for on-line work, the IRB may require consent to be provided on-line, and then have the subject respond via valid email address to verify authenticity. Similarly, for compensation or rewards or other on-line communications with the experimenter, PIN numbers may need to be generated to identify the subject on-line, and need to be kept with the security mentioned above.

Other technical details such as the sound output level and screen brightness level can not only be potential subject risk factors, but also can easily interfere with subject responses, irrespective of whether it is a psychophysical level of control desired. If one is doing multimedia experimentation, color registration might be as important as normal hearing. All of these issues require extraordinary effort on the part of the experimenter to analyze in programming for on-line presentation.

At <http://psych.hanover.edu/research/exponnet.html> many examples of questionnaire, interview, and perceptual experiments with on-line data collection are collected. The majority of those in the United States have minimal subject consent structures. [2] An example (from Canada) that displays typical consent and contact information as required by an IRB is <http://campus.upei.ca/projects/survey6/>.

The most compelling reason for on-line data collection appears to be convenience for the subject and experimenter and the large subject pools obtained. With regards to the latter, sampling theory is better met if there is a bounded sub-cultural population, such as university music majors or undergraduate liberal arts majors, than an open call to the internet world. This seems counter-intuitive. However, the steps necessary to assure veracity of subject demographics seem overwhelmingly costly with the open call. In fact, one requires a huge sample from the general internet population (which is in not the general population) to represent that population statistically. The benefits of web-based data collection are only valid if, in fact, the goals of the study are best met through the unique attributes of such an approach, and that goes far beyond simply increasing a sample size.

## NOTES

[1] However, increasing use of integrated course management systems such as Sakai and Moodle may provide a platform that can be utilized to make such student information available to experiments. IRB oversight would not be obviated in this situation.

[2] The underlying data collection protocols, server security protocols, and subject demographic data security protocols cannot be discerned.

## REFERENCES

- Honing, H., & Ladinig, O. (2008). The Potential of the Internet for Music Perception Research: A Comment on Lab-Based Versus Web-Based Studies. *Empirical Musicology Review*, Vol. 3, No. 1, pp. 4-7.
- Kerlinger, Fred N. & Lee, Howard B. (1999). *Foundations of Behavioral Research*. Wadsworth.
- McGraw, K.O., Tew, M.D., & Williams, J.E. (2000). The integrity of web-delivered experiments: Can you trust the data? *Psychological Science*, 11, 502-506.